Vector-Valued Markov Decision Processes (数理計画と決定過程論)

AUTHOR(S):

古川, 長太

数理解析研究所講究録
299巻 1977年 124-141

Vector-valued Markov Decision Processes

九大 理　古川長太

Part I. Finite-stage deterministic case

§1. 定義と準備

$\mathbb{R}^p$: p-dimensional Euclidean space

$K \subset \mathbb{R}^p$: convex cone with vertex at $0 \in \mathbb{R}^p$

ただし $K \ni \{0\}$ を仮定するが, $K \cap (-K) = \{0\}$ を仮定しない。

Def 1.1　$x, y \in \mathbb{R}^p$ に対して $x \leqq y \Leftrightarrow y - x \in K$

◎ Def 1.1 による $\leqq$ は pseudo-order である。必しも partial-order にはならない。

Def 1.2　$\Omega \subset \mathbb{R}^p$, $\Omega \neq \phi$ とする。

$x \in \Omega$ が Ω の maximal point である

$$\Leftrightarrow (\forall y \in \Omega)(x \leqq y \rightarrow y \leqq x)$$

$e(\Omega) \equiv \Omega$ の maximal point の全体

Def 1.3　$x \in \mathbb{R}^p$, $U \subset \mathbb{R}^p$ に対して

$$x + U \equiv \{ z \mid z = x + u,\ u \in U \}$$

また，$U \subset \mathbb{R}^p$, $V \subset \mathbb{R}^p$ に対して

$$U + V \equiv \{ z \mid z = u + v,\ u \in U,\ v \in V \}$$

Proposition 1.1　$x \in \mathbb{R}^p$, $U \subset \mathbb{R}^p$ に対して次のことが成立する．

$$e((x+K) \cap U) \subset e(U)$$

Corollary 1.1　$x \in U \subset \mathbb{R}^p$ に対して

$$e((x+K) \cap U) \neq \phi \Rightarrow {}^{\exists} y \in e(U) \text{ s.t. } x \leqq y$$

Proposition 1.2　A, B を 2つの parameter set とする．

$\{E_a : a \in A\}$ を，a を parameter とする $\mathbb{R}^p$ の subsets のある family

$\{F_{ab} : a \in A, b \in B\}$ を，(a, b) を parameter とする $\mathbb{R}^p$ の subsets のある family

各 $a \in A$，各 $x \in \bigcup_{b \in B} F_{ab}$ に対して

$$e\left((x+K) \cap \left(\bigcup_{b \in B} F_{ab}\right)\right) \neq \phi \quad \text{を仮定する．}$$

このとき次の関係が成立する．

$$e\left(\bigcup_{a \in A} \bigcup_{b \in B} (E_a + F_{ab})\right) = e\left(\bigcup_{a \in A} \left(E_a + e\left(\bigcup_{b \in B} F_{ab}\right)\right)\right)$$

（証明）

$$V \equiv \bigcup_{a \in A} \bigcup_{b \in B} (E_a + F_{ab}), \quad W \equiv \bigcup_{a \in A} \left(E_a + e\left(\bigcup_{b \in B} F_{ab}\right)\right)$$

$e(V) \subset e(W)$ を示す．

${}^{\forall} z \in e(V)$, $\therefore z = p + q$, $p \in E_{\hat{a}}$, $q \in F_{\hat{a}\hat{b}}$ for some $\hat{a} \in A$, some $\hat{b} \in B$

明らかに　$q \in \bigcup_{b \in B} F_{\hat{a} b}$

$q' \in \bigcup_{b\in B} F_{\hat{a}b}$ かつ $q \precsim q'$ なる q' があるとする.

$\therefore$ $q' \in F_{\hat{a}b'}$ for some $b' \in B$

$\therefore$ $(p+q') \in V$ かつ $z = p+q \precsim p+q'$

ところが $z \in e(V)$ だから $p+q' \precsim z$. $\therefore$ $q' \precsim q$

以上により $q \in e\left(\bigcup_{b\in B} F_{\hat{a}b}\right)$ が示された.

$\therefore$ $z = p+q \in E_{\hat{a}} + e\left(\bigcup_{b\in B} F_{\hat{a}b}\right) \subset W$ $\therefore$ $z \in W$

$w \in W$ を $z \precsim w$ なる点とする.

$W \subset V$ より $w \in V$

ところが $z \in e(V)$ だから $w \precsim z$

ゆえに $z \in e(W)$

次に $e(W) \subset e(V)$ を示す.

$\forall z \in e(W)$ 明らかに $z \in V$

$v \in V$, $z \precsim v$ なる v があるとする.

$\therefore$ $v = p'+q'$, $p' \in E_{a'}$, $q' \in F_{a'b'}$ for some $a' \in A$ some $b' \in B$

仮定より

$$e\left((q'+K) \cap \left(\bigcup_{b\in B} F_{a'b}\right)\right) \neq \phi$$

ゆえに Corollary 1.1 より

$\exists q'' \in e\left(\bigcup_{b\in B} F_{a'b}\right)$ s.t. $q' \precsim q''$

$\bar{v} \equiv p'+q''$

$\therefore$ $\bar{v} \in W$ かつ $v = p'+q' \precsim p'+q'' = \bar{v}$

$\precsim$ の transitive law により $z \precsim \bar{v}$

ところが $\bar{z} \in e(W)$ だから $\bar{v} \lesssim \bar{z}$

再び transitive law により $v \lesssim \bar{z}$

以上により $\bar{z} \in e(V)$ が示された □

§2. Decision process と主な結果

<u>Def 2.1</u> N-stage の deterministic vector-valued Markov decision process は 5つの要素の組 $(S, \{A_n\}_{1\le n\le N}, \{T_n\}_{1\le n\le N}, \{r_n\}_{1\le n\le N}, d)$ で表現される。ここに

S ; system の状態空間　　$s \in S$ を state と呼ぶ

A_n : $S \to \mathcal{P}(A)$ への map

ただし A はある与えられた空間で、$\mathcal{P}(A)$ は A の空でない部分集合の全体を表わす。

$A_n(s)$ を、n-th stage で利用できる action の集合、$a \in A_n(s)$ を action と呼ぶ。

$\Gamma_n \equiv \{(s, a) \mid a \in A_n(s),\ s \in S\}$

T_n ; $\Gamma_n \to S$ への map.

$\{T_n\}_{1\le n\le N}$ を system の推移法則と呼ぶ

r_n : $\Gamma_n \to \mathbb{R}^p$, n-th stage における直接利得ベクトル

d : $S \to \mathbb{R}^p$, 終端利得ベクトル

s_1 (initial state) から出発して action $a_1, a_2, \cdots, a_N$ を次々

にとり，state $s_2, s_3, \cdots, s_N$ を観測し，最後に s_{N+1} を観測して stop すると，総利得 $\sum_{n=1}^{N} r_n(s_n, a_n) + d(s_{N+1})$ を得る.

Def 2.2

$\pi = \{f_1, f_2, \cdots, f_N\}$ において，各 n につき

$f_n : S \to A$ かつ $f_n(s) \in A_n(s)$ for $\forall s \in S$

であるとき，π を (N-stage 問題に対する) policy と呼ぶ.

$\Pi \equiv$ N-stage 問題に対する policy の全体

Def 2.3 policy $\pi = \{f_1, f_2, \cdots, f_N\}$ に対して

${}^{n}\pi \equiv \{f_{n+1}, f_{n+2}, \cdots, f_N\}$ とおき，これを残りの $(N-n)$-stage に対する policy と呼ぶ. 特に ${}^{0}\pi = \pi$.

${}^{n}\Pi \equiv \{{}^{n}\pi \mid \pi \in \Pi\}$

Def 2.4

${}^{N-n}\pi = \{f_{N-n+1}, f_{N-n+2}, \cdots, f_N\}$ に対し $(1 \leq n \leq N)$

$$R^n({}^{N-n}\pi)_{s_{N-n+1}} \equiv \sum_{i=N-n+1}^{N} r_i(s_i^{\circ}, f_i(s_i^{\circ})) + d(s_{N+1}^{\circ}).$$

ただし (2.1)

$$s_i^{\circ} = T_{i-1}(s_{i-1}^{\circ}, f_{i-1}(s_{i-1}^{\circ})), \quad i = N-n+2, \cdots, N+1$$

$$s_{N-n+1}^{\circ} = s_{N-n+1}.$$

$n=0$ のときは

$$R^{0}(s_{N+1}) \equiv d(s_{N+1}) \tag{2.2}$$

Def 2.5

$$U^n(s_{N-n+1}) \equiv e\left[\bigcup_{{}^{N-n}\pi \in {}^{N-n}\Pi} R^n({}^{N-n}\pi)(s_{N-n+1})\right], \quad (1 \leq n \leq N) \tag{2.3}$$

$$U^0(s_{N+1}) \equiv R^0(s_{N+1}) \tag{2.4}$$

$U^n(s_{N-n+1})$ を簡単に $U^n(s)$ と書き、これを残りの n-stage に対する最適利得関数と呼ぶ。

Theorem 2.1　各 n，各 $s \in S$ に対し次のことを仮定する。

$$e\left[\left(R^n({}^{N-n}\hat{\pi})_{T_{N-n}(s,a)} + K\right) \cap \left(\bigcup_{{}^{N-n}\pi \in {}^{N-n}\Pi} R^n({}^{N-n}\pi)_{T_{N-n}(s,a)}\right)\right] \neq \phi$$

$$\text{for } \forall a \in A_{N-n}(s), \ \forall\, {}^{N-n}\hat{\pi} \in {}^{N-n}\Pi.$$

このとき最適利得関数の列 $\{U^n\}_{1 \leq n \leq N}$ は次の再帰式をみたす。

$$U^{n+1}(s) = e\left[\bigcup_{a \in A_{N-n}(s)} \{r_{N-n}(s,a) + U^n(T_{N-n}(s,a))\}\right] \quad \text{for } s \in S, \quad 0 \leq n \leq N-1, \tag{2.5}$$

$$U^0(s) = d(s) \qquad \text{for } s \in S. \tag{2.6}$$

(証明)

(2.6) は (2.2), (2.4) より明らか。

(2.5) が $n=0$ のとき成り立つことは (2.1), (2.3), (2.6) より明らか。

$1 \leq n \leq N$ に対して、

$$U^{n+1}(s) = e\left[\bigcup_{{}^{N-n-1}\pi} R^{n+1}({}^{N-n-1}\pi)_s\right]$$

$$= e\left[\bigcup_{{}^{N-n-1}\pi} \{r_{N-n}(s, f_{N-n}(s)) + R^n({}^{N-n}\pi)_{T_{N-n}(s, f_{N-n}(s))}\}\right]$$

$$= e\left[\bigcup_{a\in A_{N-n}(s)} \bigcup_{{}^{N-n}\pi} \{r_{N-n}(s,a) + R^n({}^{N-n}\pi)_{T_{N-n}(s,a)}\}\right]$$

上式最後の右辺に Proposition 1.2 を適用すると

$$U^{n+1}(s) = e\left[\bigcup_{a\in A_{N-n}(s)} \left\{r_{N-n}(s,a) + e\left[\bigcup_{{}^{N-n}\pi} R^n({}^{N-n}\pi)_{T_{N-n}(s,a)}\right]\right\}\right]$$

$$= e\left[\bigcup_{a\in A_{N-n}(s)} \{r_{N-n}(s,a) + U^n(T_{N-n}(s,a))\}\right]. \qquad \square$$

Def 2.6

${}^{N-n}\pi$ が残りの n-stage 問題に対し，s において optimal

$$\Leftrightarrow \quad R^n({}^{N-n}\pi)_s \in U^n(s)$$

policy $\pi \in \Pi$ が optimal $\Leftrightarrow$ $R^N(\pi)_s \in U^N(s)$ for $\forall s \in S$

Theorem 2.2. (Principle of Optimality)

$\pi^* = \{f_1^*, f_2^*, \cdots, f_N^*\}$ を optimal policy とする。このとき次のことが成立する。

$$R^n({}^{N-n}\pi^*)_{s^*_{N-n+1}} \in U^n(s^*_{N-n+1}) \qquad \text{for } \forall s_1 \in S.$$
$$n = 1, 2, \cdots, N$$

ただし

$$s_j^* = T_{j-1}(s_{j-1}^*, f_{j-1}^*(s_{j-1}^*)), \qquad j = 2, 3, \cdots, N+1$$

$$s_1^* = s_1.$$

(optimal policy π^* の final subpolicy ${}^{N-n}\pi^*$ は，残りの n-stage 問題に対して，π^*によって $(N-n)$ 番目に到達した

state s^*_{N-n+1} において optimal である。

Part II. Infinite stage stochastic case.

§3. 定義と準備

$K \subset \mathbb{R}^p$; closed convex cone with vertex at $0 \in \mathbb{R}^p$

ただし 以下では $K \cap (-K) = \{0\}$ を仮定する。

◎ Def 1.1 による $\leqq$ は partial-order である。

Def 3.1 infinite stage stochastic の vector-valued Markov decision process は 5つの組 $(S, A, (q^a_{ij}), (r^a_{ij}), \beta)$ で表現される。ここに

$S \equiv \{1, 2, 3, \cdots\}$; (countable) state space

S の元 を i, j 等で 表す。

A ; (general) action space

A の元を a で 表す。

q^a_{ij} ; action a をとることにより $i \to j$ に移る one-step transition probability

r^a_{ij} ; p 次元ベクトル値の直接利得関数

ただし $S \times S \times A$ 上で 有界 とする。

$0 < \beta < 1$; discount factor.

Def 3.2

F: $S \to A$ への map の全体からなる空間

$\pi = \{f, f, \cdots\} \equiv f^\infty \quad (f \in F)$; (stationary) policy

Π_s : (stationary) policy の全体からなる空間

$\pi \in \Pi_s$ に対し

$$I(\pi)_i \equiv E_i^\pi \left[\sum_{n=1}^\infty \beta^{n-1} r_n \right]$$

ただし r_n は n-th stage における直接利得関数を表す.
E_i^π は i を start したという条件の下で, π から induce された path の空間上の確率測度に関する条件付期待値を表す operator.

Def 3.3 $\pi^* \in \Pi_s$ にたいして

π^* が optimal $\Leftrightarrow (\forall \pi \in \Pi_s)(\forall i \in S)(I(\pi^*)_i \leq I(\pi)_i \to I(\pi)_i \leq I(\pi^*)_i)$

Def 3.4

$M(S) \equiv S$ 上の p-dim vector-valued bounded function の全体

$f \in F$ に対し $M(S) \to M(S)$ への operator T_f を次式で定義する. ; $u \in M(S)$ に対し

$$T_f u(i) = \sum_{j=1}^\infty (r_{ij}^{f(i)} + \beta u(j)) q_{ij}^{f(i)}$$

$a \in A$ に対し, $f \equiv a$ なる f により定まる T_f を T_a とかく.

Def 3.5

$u, v \in M(S)$ に対し

$u \leq v \Leftrightarrow [u(i) \leq v(i) \quad \text{for } \forall i \in S]$

Proposition 3.1

各 $f \in F$ に対し，T_f は K-monotone である．

i.e. $u \lesssim v \rightarrow T_f u \lesssim T_f v$

(証明)

▨¹ $w \in M(S)$，$w(i) \in K$ for $\forall i \in S$ なら，S 上の任意の確率測度 $p = (p_1, p_2, \cdots)$ に対して

$$\sum_{i=1}^{\infty} w(i) p_i \in K$$

∵ $$\sum_{i=1}^{\infty} w(i) p_i = \lim_{n \to \infty} \sum_{i=1}^{n} w(i) p_i$$

$$= \lim_{n \to \infty} \Big[\sum_{i=1}^{n} w(i) p_i + \underset{\substack{\uparrow \\ \mathbb{R}^p \text{の原点}}}{0} \times \Big(1 - \sum_{i=1}^{n} p_i\Big) \Big]$$

仮定により $w(1), \cdots, w(n)$, 0 はすべて K の点で K は convex より

上式の $[\cdots\cdots] \in K$.

K は closed だから $\lim_{n \to \infty} [\quad] \in K$

ゆえに $\sum_{i=1}^{\infty} w(i) p_i \in K$．ゆえに ▨¹ が示された．

さて，$u \lesssim v$，$u, v \in M(S)$ とする

$$T_f v(i) - T_f u(i) = \beta \sum_{j=1}^{\infty} (v(j) - u(j))\, q_{ij}^{f(i)}$$

$u \lesssim v$ より $v(j) - u(j) \in K$ for $\forall j$ だから ▨¹ により

$$\sum_{j=1}^{\infty} (v(j) - u(j))\, q_{ij}^{f(i)} \in K$$

K は 0 を頂点とする cone だから $\beta \Big(\sum_{j=1}^{\infty} (v(j) - u(j))\, q_{ij}^{f(i)} \Big) \in K$

∴ $T_f v(i) - T_f u(i) \in K$

これはすべての i につき成立．∴ $T_f u \lesssim T_f v$ □

Def 3.6

$\mathcal{F}(S) \equiv S$ 上で定義された，$\mathbb{R}^p$ の non-empty subset の値をとる set-valued function の全体

$U \in \mathcal{F}(S)$ に対して $\mathcal{F}(S)$ の元 $e(U)$ を次式で定義する．

$$e(U)(i) = e(U(i)) \qquad \text{for } i \in S.$$

今後，次の General Assumption を仮定する．

G. A. $\mathcal{F}(S)$ の元で恒等的に K の値をとる関数を同じ記号 K で表すことにし，

$$e\Big[\big(I(f^\infty)+K\big) \cap \Big(\bigcup_{a\in A} T_a I(f^\infty)\Big)\Big](i) \neq \Phi \quad \text{for } \forall i \in S, \ \forall f \in F$$

Proposition 3.2 G. A. は次のことと同値である．

$$\Big(\big(I(f^\infty)+K\big) \cap e\Big(\bigcup_{a\in A} T_a I(f^\infty)\Big)\Big)(i) \neq \Phi \quad \text{for } \forall i \in S, \ \forall f \in F$$

Def 3.7 $u \in M(S)$, $U \in \mathcal{F}(S)$ に対し

$$u \in U \iff [\, u(i) \in U(i) \quad \text{for } \forall i \in S \,]$$

§4 Policy improvement

Theorem 4.1

$f_0 \in F$ を任意に与え，次の iteration により $f_1, f_2, \cdots, f_n, \cdots$ を作れ．すなわち各 n につき

$$T_{f_{n+1}} I(f_n^\infty) \in \big(I(f_n^\infty)+K\big) \cap e\Big[\bigcup_{a\in A} T_a I(f_n^\infty)\Big], \quad (4.1)$$

なる $f_{n+1} \in F$ を選べ（G. A と Prop. 3.2 により (4.1) の右辺は空でないから常に可能）

このとき次のことが成立つ.

(i) $I(f_0^\infty) \leqq I(f_1^\infty) \leqq \cdots \leqq I(f_m^\infty) \leqq I(f_{m+1}^\infty) \leqq \cdots$

(ii) ある N において

$$(I(f_N^\infty)+K) \cap e\left[\bigcup_{a\in A} T_a I(f_N^\infty)\right] = \{I(f_N^\infty)\}$$

となったら, f_0 から start する (4.1) の chain では, これ以上改良は不可能で, このとき $I(f_N^\infty)$ は

$$I(f_N^\infty) \in e\left[\bigcup_{a\in A} T_a I(f_N^\infty)\right] \tag{4.2}$$

をみたす.

(証明)

(i) $T_{f_{m+1}} I(f_m^\infty) \in I(f_m^\infty) + K$ だから

$$I(f_m^\infty) \leqq T_{f_{m+1}} I(f_m^\infty)$$

T_f の K-単調より

$$I(f_m^\infty) \leqq T_{f_{m+1}} I(f_m^\infty) \leqq T_{f_{m+1}}^2 I(f_m^\infty) \leqq \cdots$$

r の有界性より $\mathbb{R}^p$ のある点, w があって

$$T_{f_{m+1}}^m I(f_m^\infty) \leqq w \quad \text{for } \forall m$$

すなわち $\{T_{f_{m+1}}^m I(f_m^\infty)\}_{m=1,2,\cdots}$ は $\leqq$ の意味で上に有界な単調列である. 仮定により K は $\mathbb{R}^p$ における regular cone だから $\{T_{f_{m+1}}^m I(f_m^\infty)\}_{m=1,2,\cdots}$ は $\mathbb{R}^p$ のある点 Q_i に i ごとに収束して, このとき $I(f_m^\infty)(i) \leqq Q_i \quad \forall i \in S$ となる.

一方, $\{T_{f_{m+1}}^m I(f_m^\infty)(i)\}_{m=1,2,\cdots}$ の p-ベクトルの成分ごとの収束を考えることにより, $Q_i = I(f_{m+1}^\infty)(i)$ が示され,

したがって　$I(f_n^\infty)(i) \lesssim I(f_{n+1}^\infty)(i) \quad \forall i$　となる.

$$\therefore I(f_n^\infty) \lesssim I(f_{n+1}^\infty)$$

(ii)　容易

§5　optimal policy の特徴づけ.

<u>Def 5.1</u>　$u \in M(S)$ に $\mathcal{F}(S)$ の元 $e(\bigcup_{a\in A} T_a u)$ を対応させる map を Φ で表す. すなわち

$$\Phi u = e(\bigcup_{a\in A} T_a u)$$

$$\therefore \Phi : M(S) \to \mathcal{F}(S)$$

<u>Def 5.2</u>　$u \in M(S)$ に対して

u が Φ の fixed point である $\Leftrightarrow$ $u \in \Phi u$

(i.e. $u(i) \in (\Phi u)(i) \quad \forall i \in S$)

<u>Def 5.3</u>　$\phi \neq \mathcal{U} \subset M(S)$ とする.

$u \in \mathcal{U}$ が $\mathcal{U}$ の maximal element である

$$\Leftrightarrow (\forall v \in \mathcal{U})(\forall i \in S)(u(i) \lesssim v(i) \to v(i) \lesssim u(i))$$

<u>Theorem 5.1</u>

$f^{*\infty}$ が optimal なら $I(f^{*\infty})$ は Φ の maximal fixed point である.

(証明)

(i)　$I(f^{*\infty})$ が Φ の fixed point であること.

i_0 を任意に fix せよ.

$$V_{i_0} \equiv \bigcup_{a \in A} T_a I(f^{*\infty})(i_0)$$

明らかに　$I(f^{*\infty})(i_0) \in V_{i_0}$

$I(f^{*\infty})(i_0) \preceq p$　なる　$p \in V_{i_0}$ があるとする.

$$\therefore\ p = T_{\hat{a}} I(f^{*\infty})(i_0) \quad \text{for some } \hat{a} \in A$$

$\hat{f}$ を
$$\hat{f}(i) = \begin{cases} \hat{a} & \text{for } i = i_0 \\ f^*(i) & \text{for } i \neq i_0 \end{cases}$$
で定義すると

$$I(f^{*\infty}) \preceq T_{\hat{f}} I(f^{*\infty})$$

かつ
$$T_{\hat{f}} I(f^{*\infty})(i_0) = T_{\hat{a}} I(f^{*\infty})(i_0) = p$$

$$\therefore\ I(f^{*\infty}) \preceq T_{\hat{f}} I(f^{*\infty}) \preceq T_{\hat{f}}^2 I(f^{*\infty}) \preceq \cdots \uparrow I(\hat{f}^{\infty}) \qquad (5.1)$$

$f^{*\infty}$ は optimal だから　(5.1) より

$$I(\hat{f}^{\infty}) \preceq I(f^{*\infty}) \qquad (5.2)$$

(5.1), (5.2) より

$$I(f^{*\infty}) = T_{\hat{f}} I(f^{*\infty})$$

ゆえに特に i_0 において

$$I(f^{*\infty})(i_0) = T_{\hat{f}} I(f^{*\infty})(i_0) = p$$

以上により　$I(f^{*\infty})(i_0) \in e(V_{i_0})$

これはすべての i_0 において成立.

$$\therefore\ I(f^{*\infty}) \in e\left[\bigcup_{a \in A} T_a I(f^{*\infty})\right]$$

$$\therefore\ I(f^{*\infty}) \in \Phi I(f^{*\infty})$$

(ii)　$I(f^{*\infty})$ が Φ の maximal fixed pt. なること.

u を Φ の任意の fixed point とする

$$\therefore \quad u(i) \in \bigcup_{a\in A} T_a u(i) \qquad \forall i$$

$$\therefore \quad u(i) = T_{a_i} u(i) \qquad \text{for some } a_i \text{, for each } i$$

$i \rightsquigarrow a_i$ なる map を $\hat{f}$ とおくと

$$u = T_{\hat{f}} u$$

$$\therefore \quad u = I(\hat{f}^\infty) \tag{5.3}$$

いま，ある i において $I(f^{*\infty})_{(i)} \lneqq u(i)$ と仮定する

(5.3) より $\quad I(f^{*\infty})_{(i)} \lneqq I(\hat{f}^\infty)(i)$

$f^{*\infty}$ は optimal だから定義により

$$I(\hat{f}^\infty)(i) \leqq I(f^{*\infty})(i)$$

$$\therefore \quad u(i) \leqq I(f^{*\infty})(i)$$

ゆえに $I(f^{*\infty})$ は Φ の maximal fixed point. $\square$

<u>Def 5.4</u>

$\tau_p \equiv M(S)$ における point-wise convergence の topology

$\Omega \equiv \{ I(f^\infty) \; ; \; f \in F \} \subset M(S)$

<u>Assumption $\mathbb{C}$</u> $\quad \Omega$ is a closed subset of $(M(S), \tau_p)$

<u>Lemma 5.1</u> $\quad$ Assumption $\mathbb{C}$ の下では Ω は Def 3.5 で導入した partial-order $\leqq$ に関する inductively ordered set である.

<u>Theorem 5.2</u> $\quad$ Assumption $\mathbb{C}$ を仮定する。このとき，$I(f^{*\infty})$ が Φ の maximal fixed point なら，$f^{*\infty}$ は optimal である.

(証明)

$I(f^{*\infty})$ が Φ の maximal fixed pt. であるとする.

ある i において

$$I(f^{*\infty})(i) \precneqq I(g^{\infty})(i) \tag{5.4}$$

なる $g \in F$ があるとする.

$$\Omega^g \rightleftharpoons \{ I(f^{\infty}) \in \Omega \,;\; I(g^{\infty}) \preccurlyeq I(f^{\infty}) \}$$

Lemma 5.1 と同様にして Ω^g は inductively ordered set であることが示される. Ω^g の $\preccurlyeq$ に関する極大元の1つを $I(\bar{g}^{\infty})$ とする.

$$\therefore\quad I(g^{\infty}) \preccurlyeq I(\bar{g}^{\infty}) \tag{5.5}$$

②: $I(\bar{g}^{\infty})$ は Φ の fixed point である.

(∵) $I(\bar{g}^{\infty})$ が Φ の fixed point でないとする.

$\therefore$ ある i_0 において

$$I(\bar{g}^{\infty})(i_0) \notin e\Big(\bigcup_{a\in A} T_a I(\bar{g}^{\infty})(i_0)\Big)$$

G.A. により

$$\exists\, p_{i_0} \in e\Big(\bigcup_{a\in A} T_a I(\bar{g}^{\infty})(i_0)\Big) \quad \text{s.t.}\ I(\bar{g}^{\infty})(i_0) \preccurlyeq p_{i_0}$$

明らかに $p_{i_0} \neq I(\bar{g}^{\infty})(i_0)$ で,

$$p_{i_0} = T_{a_{i_0}} I(\bar{g}^{\infty})(i_0) \qquad \text{for some } a_{i_0} \in A \tag{5.6}$$

$$\bar{S} \rightleftharpoons \{ i \in S \,;\; I(\bar{g}^{\infty})_i \notin e\Big(\bigcup_{a\in A} T_a I(\bar{g}^{\infty})(i)\Big) \}$$

$i_0 \in \bar{S}$ の各 i_0 に対して (5.6) のように a_{i_0} を定めよ.

$\hat{f}$ を次のように定める.

$$\hat{f}(i) = \begin{cases} a_i & \text{for } i \in \bar{S} \\ \bar{g}(i) & \text{for } i \notin \bar{S} \end{cases}$$

$$\therefore \quad I(\bar{g}^\infty)(i) \preceq T_{\hat{f}} I(\bar{g}^\infty)(i) \quad \forall i$$

$$\therefore \quad I(\bar{g}^\infty) \preceq T_{\hat{f}} I(\bar{g}^\infty) \tag{5.7}$$

$i_0 \in \bar{S}$ においては $p_{i_0} = T_{a_{i_0}} I(\bar{g}^\infty)(i_0) \neq I(\bar{g}^\infty)(i_0)$ だから

$$I(\bar{g}^\infty) \neq T_{\hat{f}} I(\bar{g}^\infty) \tag{5.8}$$

(5.7) より帰納的に

$$I(\bar{g}^\infty) \preceq T_{\hat{f}} I(\bar{g}^\infty) \preceq T_{\hat{f}}^2 I(\bar{g}^\infty) \preceq \cdots \uparrow I(\hat{f}^\infty) \tag{5.9}$$

しかるに $I(\bar{g}^\infty)$ は $\preceq$ に関する極大元 だから (5.9) より

$$I(\bar{g}^\infty) = I(\hat{f}^\infty)$$

ゆえに (5.9) より $I(\bar{g}^\infty) = T_{\hat{f}} I(\bar{g}^\infty)$ となりこれは (5.8) に反する. 以上により $I(\bar{g}^\infty)$ は Φ の fixed point である.

次に (5.4) でとられた i において (5.4) (5.5) より

$$I(f^{*\infty})(i) \preceq I(g^\infty)(i) \preceq I(\bar{g}^\infty)(i) \tag{5.10}$$

ところが $I(\bar{g}^\infty)$ は Φ の fixed point であり, 仮定により $I(f^{*\infty})$ は Φ の maximal fixed point だから (5.10) より

$$I(f^{*\infty})(i) = I(\bar{g}^\infty)(i)$$

ゆえに $f^{*\infty}$ は optimal $\square$

Theorem 5.3

(i) A : compact metric space

(ii) 各iに対し，r_{ij}^{a} は j につき一様に a に関して連続

(iii) 各iに対し，$\{q_{ij}^{a}\}_{j=1,2,\cdots}$ は S 上の probability measure として weak topology の意味で a に関して連続

$\Longrightarrow$

Assumption $\mathbb{C}$ 成立.

(証明は長くなるので省略する)

Note.1 S, A がともに finite set なら Thm 5.3 における (i),(ii),(iii) 及び general assumption G.A. はすべてみたされる. すなわち本論文における必要な仮定はすべてみたされる.

Note 2 Thm 4.1 における (4.2) は, $I(f_N^{\infty})$ が Φ の fixed point であることを示している.